



**UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office**

Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.
08/746,981	11/19/96	LEE	087.35020X00

020457 LM61/0318
ANTONELLI TERRY STOUT KRAUS
1300 NORTH SEVENTEENTH STREET SUITE 1800
ARLINGTON VA 22209

EXAMINER
STORM, D

ART UNIT	PAPER NUMBER
2741	4

DATE MAILED: 03/18/98

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner of Patents and Trademarks

Office Action Summary

Application No.
08/746,981

Applicant(s)
Lee et al.

Examiner
Donald L. Storm

Group Art Unit
2741



☒ Responsive to communication(s) filed on Nov 19, 1996

☐ This action is **FINAL**.

☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11; 453 O.G. 213.

A shortened statutory period for response to this action is set to expire 3 month(s), or thirty days, whichever is longer, from the mailing date of this communication. Failure to respond within the period for response will cause the application to become abandoned. (35 U.S.C. § 133). Extensions of time may be obtained under the provisions of 37 CFR 1.136(a).

Disposition of Claims

☒ Claim(s) 1-3 is/are pending in the application.

Of the above, claim(s) _____ is/are withdrawn from consideration.

☐ Claim(s) _____ is/are allowed.

☒ Claim(s) 1-3 is/are rejected.

☐ Claim(s) _____ is/are objected to.

☐ Claims _____ are subject to restriction or election requirement.

Application Papers

☒ See the attached Notice of Draftsperson's Patent Drawing Review, PTO-948.

☒ The drawing(s) filed on Nov 19, 1996 is/are objected to by the Examiner.

☐ The proposed drawing correction, filed on _____ is ☐ approved ☐ disapproved.

☒ The specification is objected to by the Examiner.

☒ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. § 119

☒ Acknowledgement is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d).

☒ All ☐ Some* ☐ None of the CERTIFIED copies of the priority documents have been

☒ received.

☐ received in Application No. (Series Code/Serial Number) _____.

☐ received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

*Certified copies not received: _____

☐ Acknowledgement is made of a claim for domestic priority under 35 U.S.C. § 119(e).

Attachment(s)

☒ Notice of References Cited, PTO-892

☒ Information Disclosure Statement(s), PTO-1449, Paper No(s). 2

☐ Interview Summary, PTO-413

☒ Notice of Draftsperson's Patent Drawing Review, PTO-948

☐ Notice of Informal Patent Application, PTO-152

--- SEE OFFICE ACTION ON THE FOLLOWING PAGES ---

DETAILED ACTION

1. The Examiner suggests that claiming the two nodes in the output layer disclosed on page 5 at lines 19-21 and Fig. 2, as a whole with other limitations, may patentably distinguish the invention from prior art. The prior art currently of record does not anticipate or make obvious the disclosed phoneme segmenter with a node that produces OUT(0) and another node that produces OUT (1). As currently claimed, OUT(0) and OUT (1) are names that distinguish a first value of output from a second value of output, so that different values are identified. As claimed, the MLP output that produces the two values may be one output that produces different values at different times. The claims currently do not recite a first output node or first output and a second output node or second output.

Priority

2. Receipt is acknowledged of papers submitted under 35 U.S.C. 119(a)-(d), which papers have been placed of record in the file.

Oath/Declaration

3. The oath or declaration is defective. A new oath or declaration in compliance with 37 CFR 1.67(a) identifying this application by application number and filing date is required. See MPEP §§ 602.01 and 602.02.

The oath or declaration is defective because:

It does not identify the city and state or foreign country of residence of each inventor as required by 37 CFR § 1.63 and § 1.33(a).

Drawings

4. The drawings are objected to under 37 CFR 1.83(a) because they fail to show the claimed invention. The drawings must show every feature of the invention specified in the claims. Therefore, all the features of all the claims must be shown or the feature(s) canceled from the claim(s). See MPEP § 608.02(d). Correction is required. At a minimum, representation of the following features should be added to the drawings:

“two frames preceding the final frame” (claim 1). The Examiner suggests that Item 23 could be modified to show this feature, for example, --within 2 frames of final?--.

Specification

5. The disclosure is objected to because of the following informalities. Appropriate correction is required.

On page 12, before line 1, the first (unnumbered) equation ends with “n-”, as though cut off at the margin during printing. If “n-” is correct, please define “n-” near this equation. The symbol $ID(j)$ is not defined. The symbol $IN(i)$ is not defined, although the text defines $IN(j)$, which is not in the equations.

On page 12, before line 1, in the second (unnumbered) equation, the symbol $ID(N-1)$ is not defined.

On page 12, before line 1, in the third (unnumbered) equation, the symbol $UT(k)$ is not defined. It should probably be $OUT(k)$, which is defined in the text.

Claim Rejections - 35 USC § 112

6. Claims 1-3 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.
7. Regarding claim 1, the phrase "converted from voice made" (page 14, lines 4-5) needs clarification. Perhaps a word or phrase is missing.
8. Regarding claim 1, the phrase "the weight" obtained through learning (page 14, line 24 and again, page 15, line 3) lacks definite antecedence in the claim. Although weights were recited as initialized, no single, particular weight was set forth. To search and apply prior art, the Examiner has interpreted "the weight" as --the weights--.
9. Regarding claim 1, the phrase "said characteristic vectors" (page 14, line 23 and again, page 15, line 4) lacks definite antecedent basis in the claims. Both characteristic vectors by vocal frames and an inter-frame characteristic vector were previously recited. As recited in the claim, it would not be clear to an artisan if the inter-frame vector, the characteristic vectors by frames, or both define the scope of the claim in this inputting step. The specification (page 9, line 7-page 10, line 4) teaches inter-frame characteristic vectors as input to the MLP. Note the plural inter-frame vectors taught in the specification vice the singular, recited in the claim. Is it one inter-frame vector for each current analyzed frame?

10. Regarding claim 1, the phrase “the current analyzed frame” (page 15, lines 6-7) lacks antecedent basis in the claim. To search and apply prior art, the Examiner has assumed that there is an iterative analysis procedure underway that treats each frame sequentially, so that a particular frame is being evaluated at any iteration; that frame may be called “the current frame.” The Examiner suggests that antecedence could be provided earlier in the claim by a phrase such as --receiving said characteristic vector [or is it vectors?] of a frame for current analysis--.

11. Regarding claim 1, the phrase “the final frame” of incoming voice (page 15, lines 7-8) lacks antecedent basis in the claim. An artisan would not be reasonably apprised of the scope of the claim because “incoming voice” is not a claimed element, so an artisan would not know to what portion the final frame is final. To search and apply prior art, the Examiner has interpreted this phrase to describe some endpoint in the vocal sample that was somehow detected, marked, and presented to the MLP.

12. Claims 2 and 3 inherit the problems of claim 1 by dependency.

13. Regarding claim 2, the phrase “the overall incoming vocal samples” (page 15, lines 12-13) lacks antecedent basis. An artisan would not be reasonably apprised of the scope of the claim because the claim provides no basis on which to determine what limitation is provided by the recitation of “framing . . . with respect to the overall incoming vocal samples”. To search and apply prior art, the Examiner has interpreted this limitation as framing the overall length of each sequential segment of digitized voice samples as recited on page 14 at lines 14-15.

Claim Rejections - 35 USC § 102

14. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless --

(a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

15. Claims 1-2 are rejected under 35 U.S.C. 102(a) as being anticipated by Youngjoo Soo and Youngjik Lee, "Phoneme Segmentation of Continuous Speech Using Multi-Layer Perceptron," Proceedings of Fourth International Conference on Spoken Language Processing ICSLP '96 (Philadelphia), 3-6 Oct. 1996, pp.1297-1300.

16. Regarding claim 1, Soo and Lee discloses a phoneme segmentation method that teaches:

voice input converted to digital samples as a read speech data base, sampled at 16 kHz with 16-bit resolution (page 1298, column 2, lines 33-37);

extracting a characteristic vector suitable for phoneme division from the vocal sample input as the primary features and final features for phoneme segmenter input, the inter-frame features that are differences between two adjacent frames (page 1298, column 1, lines 10 ff. and lines 36-41);

a multilevel neural network in the form of a multi-layer perceptron (MLP) phoneme dividing portion that finds and outputs the border of phoneme using the characteristic vector as a MLP-based phoneme segmenter (page 1298, column 1, lines 44 ff.) that decides the positions of phoneme boundaries using the output value of the MLP (page 1298, column 1, lines 53-54);

outputting position information on the border of phoneme of the MLP phoneme dividing portion in the form of frame position as the detected boundaries (page 1299, column 1, line 13);

sequentially segmenting and framing the digitized voice samples as phoneme segmentation and framing in 16 msec frames with a Hamming window at a shift rate of 10 msec (page 1298, column 1, lines 3-7);

extracting characteristic vectors by vocal frames as the primary features (page 1298, column 1, lines 10 ff.);

extracting an inter-frame characteristic vector of the difference between nearby frames of the characteristic vectors by frames as final features for phoneme segmenter input, the inter-frame features that are differences between two adjacent frames (page 1298, column 1, lines 36-41);

normalizing the maximum and minimum of the inter-frame characteristics as normalizing to lie between -1 and +1 (page 1298, column 1, line 42);

initializing weights present between an input layer and hidden layer and between the hidden layer and output layer of the MLP as the initial weight values (page 1298, column 2, lines 13-14) for the MLP with one hidden layer and an output layer (page 1298, column 1, lines 44-45);

designating an output target data of the MLP for learning as target data with a value of +1 at the phoneme boundary and -1 or similar value in other position (page 1298, column 2, lines 9-10);

storing and finishing information in the weight[s] obtained through learning as learning all cases of speech input patterns (page 1298, column 2, line 12);

the standard of the MLP if the reduction rate of mean squared error converges within a permissible limit as the MLP-based phoneme segmenter (page 1298, column 1, lines 44 ff.) trained with the learning rate set to 0.0005 (page 1298, column 2, lines 12-13) for the minimum mean-squared error criterion (page 1298, column 2, lines 6-7);

reading the weight[s] obtained as the capability of the trained MLP of detecting these [phoneme] boundaries (page 1297, column 2, lines 28-29);

receiving the characteristic vectors as introducing well-defined features to make patterns representing the boundary and nonboundary of phonemes (page 1297, column 2, lines 26-28);

performing an operation of phoneme border discrimination as detecting these boundaries and segmenting continuous speech into their corresponding phonemes (page 1298, column 2, lines 29-30);

generating an output value as the output value of the MLP (page 1298, column 1, line 54);

discriminating the phoneme border according to the output value the segmentation decision made by the threshold value of the MLP output (page 1299, column 2, lines 45-46);
and

if the current analyzed frame arrives two frames preceding the final frame of incoming voice, outputting a frame number indicative of the border of phoneme as a final result as regarding the position of the third frame as a phoneme boundary (page 1298, column 1, lines 55-57) because the transition takes at least a couple of frames in normal cases (page 1299, column 1, lines 23-25).

17. Regarding claim 2, Soo and Lee teaches the limitations as for claim 1 and further teaches:

frames formed by a Hamming window in a length 16 msec every 10 msec, with respect to the overall incoming vocal samples the same as teaching the sequentially segmenting and framing element of claim 1.

Claim Rejections - 35 USC § 103

18. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(f) or (g) prior art under 35 U.S.C. 103(a).

19. Claim 3 is rejected under 35 U.S.C. 103(a) as being unpatentable over Youngjoo Soo and Youngjik Lee, "Phoneme Segmentation of Continuous Speech Using Multi-Layer Perceptron," Proceedings of Fourth International Conference on Spoken Language Processing ICSLP '96 (Philadelphia), 3-6 Oct. 1996, pp.1297-1300.

20. Regarding claim 3, Soo and Lee teaches the limitations as for claim 1 and further teaches: determining that if output value OUT(0) is positive, an analyzed frame is the border of phonemes as when the output of the MLP is greater than the threshold value, regarding the third frame as a phoneme boundary (page 1299, column 1, lines 54-57);

Soo and Lee, however, does not explicitly teach determining if output value OUT(1) is positive, the frame is not the border of phonemes.

Because Soo and Lee teach a positive OUT(0) when the output of the MLP is greater than the threshold value to be indicative of the phoneme boundary, the converse would have been obvious to one of ordinary skill in the art of speech recognition, namely, that a positive output value OUT(1) when the threshold is greater than the output of the MLP would determine that the frame is not the border of phonemes.

Conclusion

21. The Examiner refers the Applicant to MPEP §706.02(b) where actions are listed that are available to overcome rejection based on 35 U.S.C. 102(a):

(1) Persuasively arguing that the claims are patentably distinguishable from the prior art;

(2) Amending the claims to patentably distinguish over the prior art;

(3) Filing an affidavit or declaration under 37 CFR 1.131. See MPEP § 715 for information on the requirements of 37 CFR 1.131 affidavits.

(4) Filing an affidavit or declaration under 37 CFR 1.132 showing that the reference invention is not by "another." See MPEP § 715.01(a), § 715.01(c), and § 716.10;

- (5) Perfecting a claim to priority under 35 U.S.C. 119(a) - (d);
- (6) Perfecting priority under 35 U.S.C. 119(e) by amending the specification of the application to contain a specific reference to a provisional application in accordance with 37 CFR 1.78(a)(4).

22. The following references made of record and not relied upon are considered pertinent to applicant's disclosure:

Lennig et al. [U. S. Patent Number 4,956,865] teaches word endpoint determination and recognition suitable for isolated or continuous speech using signal features and their differences by frame, and setting two frames from the end frames as the threshold for determining word-end frames.

Hong C. Leung, Benjamin Chigier, and James R. Glass, "A Comparative Study of Signal Representations and Classification Techniques for Speech Recognition," IEEE, 1992, pp. 680-682, teaches phoneme boundary detection by an MLP with one hidden layer using frame-based differences in normalized spectrum/cepstrum characteristics to calculate a probability that each frame is the boundary of a phoneme.

Chigier [U. S. Patent Number 5,638,487] teaches identifying sub-word boundary frames, such as phonemes, using an MLP with two nodes in the output layer based on cepstral coefficients by comparing probabilities for each frame to two thresholds.

Don R. Hush and Bill G. Horne, "Progress in Supervised Neural Networks," IEEE Signal Processing Magazine, January 1993, pp. 8-39, teaches training, finishing, and preparing a MLP for use.

Watari et al. [U. S. Patent Number 4,592,085] teaches phoneme boundary determination by minimum-distances comparison of features in a time frame to characteristic features.

Dan Hammerstrom, "Working With Neural Networks," IEEE Spectrum, July 1993, pp.46-53, teaches the basic steps of training and using neural networks, such as storing and retrieving the weights, the size, and the connections of the network.

23. Any response to this action should be mailed to:

Commissioner of Patents and Trademarks
Washington, D.C. 20231

or faxed to:

(703) 308-9051, (for formal communications intended for entry)

Or:

(703) 305-9508, (for informal or draft communications, and please label "PROPOSED" or "DRAFT")

Hand-delivered responses should be brought to Crystal Park II, 2121 Crystal Drive, Arlington, VA., Sixth Floor (Receptionist).

24. ***** **IMPORTANT NOTICE** *****

Effective November 16, 1997, the Examiner handling this application will be assigned to a new Unit as a result of the consolidation into Technology Center 2700. See the forth coming Official Gazette notice dated November 11, 1997. For any written or facsimile communication submitted ON OR AFTER November 16, 1997, this Examiner, who was assigned to Art Unit 2308, will be assigned to **ART UNIT 2741**. Please include the new Art Unit in the caption or

heading of any communication submitted after the November 16, 1997 date. Your cooperation in this matter will assist in the timely processing of the submission and is appreciated by the Office.

25. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Donald L. Storm whose telephone number is (703)305-3941. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth, can be reached at (703)308-4825. Any inquiry of a general nature or relating to the status of this application should be directed to the Group Receptionist whose telephone number is (703)305-3900.



DAVID R. HUDSPETH
SUPERVISORY PATENT EXAMINER
GROUP 2700